Radboud Repository

Radboud University Nijmegen

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/22748

Please be advised that this information was generated on 2018-07-07 and may be subject to change.

© 1996 Oxford University Press

Human Molecular Genetics, 1996, Vol. 5, No. 7 1035–1041

Positional cloning of the gene for X-linked retinitis pigmentosa 3: homology with the guanine-nucleotide-exchange factor RCC1

Ronald Roepman[1], Gerard van Duijnhoven[1], Thomas Rosenberg[2], Alfred J. L. G. Pinckers[3], Liesbeth M. Bleeker-Wagemakers[4], Arthur A. B. Bergen[4], Jan Post[5], Alfred Beck[6], Richard Reinhardt[6], Hans-Hilger Ropers[1,6], Frans P. M. Cremers[1] and Wolfgang Berger[6,*]

Departments of [1]Human Genetics and [3]Ophthalmology, University Hospital Nijmegen, Nijmegen, The Netherlands, [2]National Eye Clinic for the Visually Impaired, Hellerup, Denmark, [4]Department of Ophthalmogenetics, Ophthalmic Research Institute, Amsterdam, The Netherlands, [5]Clinical Genetics Center, Utrecht, The Netherlands and [6]Max-Planck-Institut für Molekulare Genetik, Berlin, Germany

Received April 29, 1996; Accepted May 2, 1996

EMBL accession no. X97668

The gene for retinitis pigmentosa 3 (*RP3*), the most frequent form of X-linked RP (XLRP), has been mapped previously to a chromosome interval of less than 1000 kbp between the DXS1110 marker and the OTC locus at Xp21.1-p11.4. Employing a novel technique, 'YAC Representation Hybridization (YRH)', we have recently identified a small XLRP associated microdeletion in this interval, as well as several putative exons including the 3′ end of a gene that was truncated by the deletion. cDNA library screening and sequencing of a cosmid centromeric to the deletion has now enabled us to identify numerous additional exons and to detect several point mutations in patients with XLRP. The predicted gene product shows homology to RCC1, the guanine-nucleotide-exchange factor (GEF) of the Ras-like GTPase Ran. Our findings suggest that we have cloned the long-sought *RP3* gene, and that it may encode the GEF of a retina-specific GTP-binding protein.

INTRODUCTION

Retinitis pigmentosa (RP), a heterogeneous group of disorders characterized by night blindness, progressive constriction of the visual field and 'pigmented' fundus abnormalities, is due to the gradual degeneration of photoreceptor cells which begins in the periphery and progresses to the central region of the retina. In advanced stages, patients present with tunnel vision and may become legally blind. About 80% of the patients have autosomal recessive forms of RP. Autosomal dominant and the more severe X-linked RP (XLRP) account each for approximately 10% of the cases, but in Denmark and the UK, a much higher proportion of XLRP has been reported (1,2).

At least three different gene loci are involved in X-linked retinitis pigmentosa (XLRP), which have been mapped by linkage studies and heterogeneity testing to Xp11.2-p11.4 (*RP2*), Xp11.4-p21.1 (*RP3*) and farther distal on Xp (*RP6*), respectively (3–5). By molecular characterization of two deletions in patients with RP and other X-linked disorders, the gene for the most frequent form of XLRP, retinitis pigmentosa 3 (*RP3*), has been assigned previously to an interval of less than 1000 kilobases (kb) between the DXS1110 marker and the ornithine transcarbamylase (OTC) gene (6).

In previous studies by our group that led to the molecular elucidation of two other eye disorders (7,8), intragenic microdeletions of a few kbp had been instrumental for fine mapping and cloning of the respective genes. With conventional (Southern blotting and PCR based probe screening) techniques, such small microdeletions are difficult to find, however, even in genomic intervals of 1000 kb or less. To overcome this difficulty, we have recently developed a novel 'YAC Representation Hybridisation' technique which was employed for microdeletion screening in the chromosomal interval carrying the *RP3* gene (9). In one out of 30 patients with XLRP, a 6.4 kb microdeletion was detected. By sequencing of a cosmid spanning this deletion and by cDNA library screening, several exons were found, including the 3′ end of a gene that was partially deleted in this patient. None of these exons contained an open reading frame, however, and the search for point mutations in patients with XLRP was unsuccessful.

Therefore, in order to identify additional exons from the 5′ end of this gene, we have now isolated and sequenced a neighboring cosmid from a previously established cosmid contig, as well as two *Eco*RI fragments separating this cosmid from the (more distal) one that had been sequenced previously. Computer-assisted analysis of this sequence predicted numerous additional exons which could be confirmed by cDNA cloning and sequencing, and in several patients with XLRP, small mutations were detected which were not observed in healthy controls. Sequence comparisons have revealed that the deduced product of this gene shows strong structural similarities with RCC1, the guanine nucleotide exchange factor of the Ras-like GTPase Ran. These findings suggest that we have identified the long-sought *RP3* gene, and

*To whom correspondence should be addressed

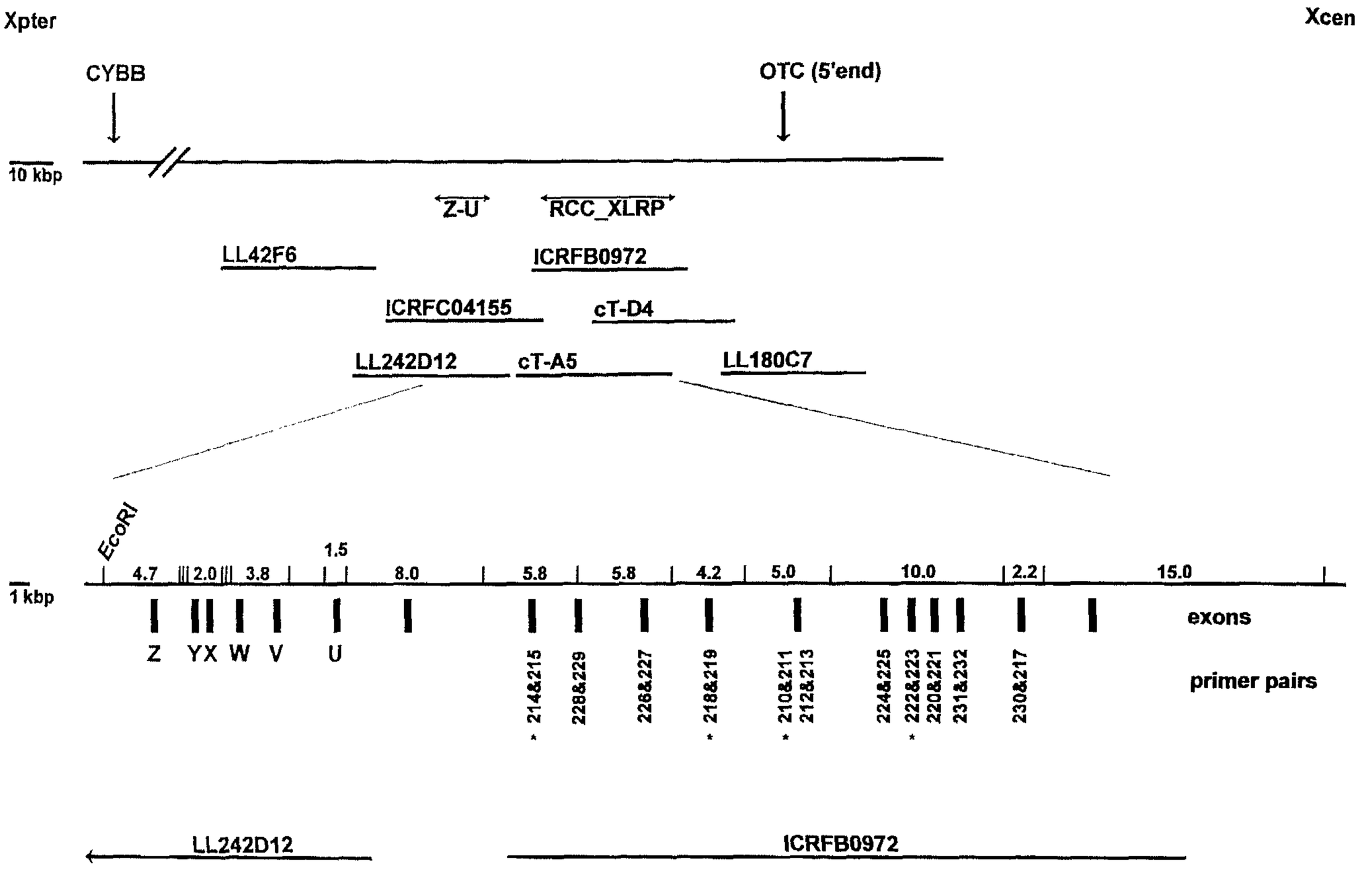

Figure 1. Physical map of the region containing the *RP3* gene. The upper part of the figure shows the cosmid contig covering the region between the 3′ end of the *RP3* gene and the 5′ end of the OTC gene. Cosmids derived from the ICRF and Lawrence Livermore (LL) libraries were published previously (9). Clones cT-A5 and -D4 were isolated from a library generated by subcloning a YAC (see Materials and Methods). The lower part of this figure displays the *Eco*RI restriction map and exon distribution in a 70 kbp region harboring the *RP3* gene. Cosmid clones LL242D12 and ICRFB0972 were sequenced. The sequence contig was completed by sequencing the 8.0 and 5.8 kbp fragments between them. Exons were identified by aligning cDNA and genomic sequences. Fragments U-V were described previously (9), and all other exon fragments correspond to cDNAs which were isolated by screening a retina library with cosmid clone cT-A5. Primer numbers given for 10 of the novel exons refer to Table 1. The asterisks indicate fragments containing sequence alterations confined to patients.

they may provide a clue to its role in the pathogenesis of retinitis pigmentosa.

RESULTS

Identification of protein coding sequences from the *RP3* critical region

In order to extend the cDNA of a previously identified candidate gene for *RP3* (9) in the 5′ direction, we have sequenced the neighboring cosmid clone ICRFB0972 (see Fig. 1) by a shotgun approach. Sequence data were assembled in 15 contigs, covering 39.776 bp proximal to the cosmid LL242D12 (Fig. 1), which was sequenced recently (9). Moreover, the gap between these two cosmids was closed by sequencing two intervening 8.0 and 5.8 kbp *Eco*RI fragments by primer walking. In total, we have sequenced approximately 75 kbp from the genomic region carrying the *RP3* gene. In parallel another cosmid, cT-A5, was employed to screen a cDNA library from adult retina. This cosmid clone which is almost identical to cosmid ICRFB0972 was derived from a cosmid library of YAC ICRFy900E0701 (see Fig. 1). The cDNA screening revealed six positive clones, four of which were characterized in more detail by Southern blot hybridization and sequence analysis. Seven partially overlapping cDNA clones were assembled by using the program GELMERGE (10). The resulting cDNA consensus sequence was aligned to genomic sequences from cosmids ICRFB0972 and LL242D12 and the 8.0 and 5.8 kbp *Eco*RI fragments. In this way 12 additional exons were identified (Figs 1 and 2). One hundred and one nucleotides at the 3′ end of the consensus cDNA sequence did not match genomic sequences which points to a gap in our genomic sequence contig. Since the penultimate exon of the cDNA is present on the 5.8 kbp *Eco*RI fragment, this gap is most likely between this and the adjacent 8.0 kbp *Eco*RI fragment. Five of the exons were also identified by the GRAIL exon prediction program (releases 2.0 and 1.3, see Materials and Methods).

Northern blot hybridization with a probe containing 1177 bp from the 5′ end of the cDNA detected a transcript of approximately 3 kbp in heart, brain, placenta, lung, liver, skeletal muscle, and pancreas (Fig. 3). A FASTA search with the consensus cDNA sequence of 1557 bp failed to detect significant homologies with GENEMBL entries. It contains a contiguous reading frame of 518 amino acid residues (Fig. 2), which was used to perform database searches by employing the BLASTP as well as the FASTA programs (10). Significant homologies were detected with RCC1 proteins from several species, including human, hamster, *Xenopus*, *Drosophila*, and yeast. The most significant homology was

```
1    CCGCATGGCGACCCTCCGGAGATCAACTACAACCGCGCTCCCGGAAGAGCTGATGCCCGA   60
      R  M  A  T  L  R  R  S  T  T  T  A  L  P  E  E  L  M  P  D

61   TTCGGGTGCTGTGTTTACATTTGGGAAAAGTAAATTTGCTGAAAATAATCCCGGTAAATT  120
      S  G  A  V  F  T  F  G  K  S  K  F  A  E  N  N  P  G  K  F

121  CTGGTTTAAAAATGATGTCCCTGTACATCTTTCATGTGGAGATGAACATTCTGCTGTTGT  180
      W  F  K  N  D  V  P  V  H  L  S  C  G  D  E  H  S  A  V  V

181  TACCGGAAATAATAAACTTTACATGTTTGGCAGTAACAACTGGGGTCAGTTAGGATTAGG  240
      T  G  N  N  K  L  Y  M  F  G  S  N  N  W  G  Q  L  G  L  G

241  ATCAAAGTCAGCCATCAGCAAGCCAACATGTGTCAAAGCTCTAAAACCTGAAAAAGTGAA  300
      S  K  S  A  I  S  K  P  T  C  V  K  A  L  K  P  E  K  V  K

301  ATTAGCTGCCTGTGGAAGGAACCACACCCTGGTGTCAACAGAAGGAGGCAATGTATATGC  360
      L  A  A  C  G  R  N  H  T  L  V  S  T  E  G  G  N  V  Y  A

361  AACTGGTGGAAATAATGAAGGACAGTTGGGGCTTGGTGACACCGAAGAAAGAAACACTTT  420
      T  G  G  N  N  E  G  Q  L  G  L  G  D  T  E  E  R  N  T  F

421  TCATGTAATTAGCTTTTTTACATCCGAGCATAAGATTAAGCAGCTGTCTGCTGGATCTAA  480
      H  V  I  S  F  F  T  S  E  H  K  I  K  Q  L  S  A  G  S  N

481  TACTTCAGCTGCCCTAACTGAGGATGGAAGACTTTTTATGTGGGGTGACAATTCCGAAGG  540
      T  S  A  A  L  T  E  D  G  R  L  F  M  W  G  D  N  S  E  G

541  GCAAATTGGTTTAAAAAATGTAAGTAATGTCTGTGTCCCTCAGCAAGTGACCATTGGGAA  600
      Q  I  G  L  K  N  V  S  N  V  C  V  P  Q  Q  V  T  I  G  K

601  ACCTGTCTCCTGGATCTCTTGTGGATATTACCATTCAGCTTTTGTAACAACAGATGGTGA  660
      P  V  S  W  I  S  C  G  Y  Y  H  S  A  F  V  T  T  D  G  E

661  GCTATATGTGTTTGGAGAACCTGAGAATGGGAAGTTAGGTCTTCCCAATCAGCTCCTGGG  720
      L  Y  V  F  G  E  P  E  N  G  K  L  G  L  P  N  Q  L  L  G

721  CAATCACAGAACACCCCAGCTGGTGTCTGAAATTCCGGAGAAGGTGATCCAAGTAGCCTG  780
      N  H  R  T  P  Q  L  V  S  E  I  P  E  K  V  I  Q  V  A  C

781  TGGTGGAGAGCATACTGTGGTTCTCACGGAGAATGCTGTGTATACCTTTGGGCTGGGACA  840
      G  G  E  H  T  V  V  L  T  E  N  A  V  Y  T  F  G  L  G  Q

841  ATTTGGTCAGCTGGGTCTTGGCACTTTTCTTTTTGAAACTTCAGAACCCAAAGTCATTGA  900
      F  G  Q  L  G  L  G  T  F  L  F  E  T  S  E  P  K  V  I  E

901  GAATATTAGGGATCAAACAATAAGTTATATTTCTTGTGGAGAAAATCACACAGCTTTGAT  960
      N  I  R  D  Q  T  I  S  Y  I  S  C  G  E  N  H  T  A  L  I

961  AACAGATATCGGCCTTATGTATACTTTTGGAGATGGTCGCCACGGAAAATTAGGACTTGG 1020
      T  D  I  G  L  M  Y  T  F  G  D  G  R  H  G  K  L  G  L  G

1021 ACTGGAGAATTTTACCAATCACTTCATTCCTACTTTGTGCTCTAATTTTTTGAGGTTTAT 1080
      L  E  N  F  T  N  H  F  I  P  T  L  C  S  N  F  L  R  F  I

1081 AGTTAAATTGGTTGCTTGTGGTGGATGTCACATGGTAGTTTTTGCTGCTCCTCATCGTGG 1140
      V  K  L  V  A  C  G  G  C  H  M  V  V  F  A  A  P  H  R  G

1141 TGTGGCAAAAGAAATTGAATTCGATGAAATAAATGATACTTGCTTATCTGTGGCGACTTT 1200
      V  A  K  E  I  E  F  D  E  I  N  D  T  C  L  S  V  A  T  F

1201 TCTGCCGTATAGCAGTTTAACCTCAGGAAATGTACTGCAGAGGACTCTATCAGCACGTAT 1260
      L  P  Y  S  S  L  T  S  G  N  V  L  Q  R  T  L  S  A  R  M

1261 GCGGCGAAGAGAGAGGGAGAGGTCTCCAGATTCTTTTTCAATGAGGAGAACACTACCTCC 1320
      R  R  R  E  R  E  R  S  P  D  S  F  S  M  R  R  T  L  P  P

1321 AATAGAAGGGACTCTTGGCCTTTCTGCTTGTTTTCTCCCCAATTCAGTCTTTCCACGATG 1380
      I  E  G  T  L  G  L  S  A  C  F  L  P  N  S  V  F  P  R  C

1381 TTCTGAGAGAAACCTCCAAGAGAGTGTCTTATCTGAACAGGACCTCATGCAGCCAGAGGA 1440
      S  E  R  N  L  Q  E  S  V  L  S  E  Q  D  L  M  Q  P  E  E

1441 ACCAGATTATTTGCTAGATGAAATGACCAAAGAAGCAGAGATAGATAATTCTTCAACTGT 1500
      P  D  Y  L  L  D  E  M  T  K  E  A  E  I  D  N  S  S  T  V

1501 AGAAAGCCTTGGAGAAACTACTGATATCTTAAACATGACACACATCATGAGCCTGAA    1557
      E  S  L  G  E  T  T  D  I  L  N  M  T  H  I  M  S  L
```

Figure 2. cDNA consensus sequence with the deduced protein translation. The boundaries between the different exons are represented by arrowheads. The sequence contains a reading frame, encoding 518 amino acid residues. A translation start site was not identified, indicating an incomplete 5′ end.

observed with human, hamster and *Xenopus* RCC1 proteins, i.e. approximately 30% identity within an overlap of 450 amino acid residues. Amino acid identity with the *Drosophila* and yeast RCC1 peptides was 26 and 20%, respectively.

In a previous study (9), we had identified the 3′ end of a gene, which was disrupted in one patient with XLRP by a 6.4 kbp

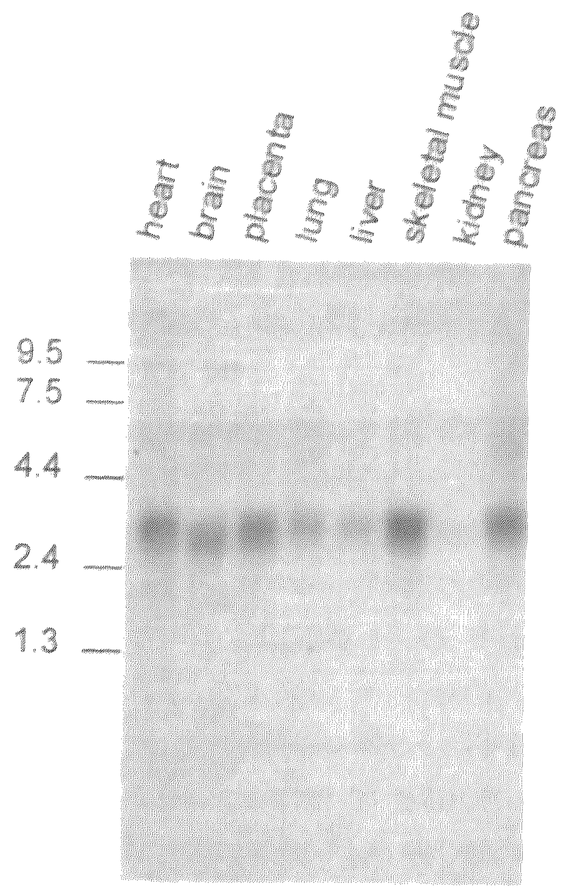

Figure 3. Autoradiograph of a Northern blot containing poly(A)$^+$-RNA from various adult tissues after hybridization with a cDNA probe containing 1177 bp from the 5′ end of the cDNA presented in Figure 2. Sizes of the marker fragments are given in kbp. The size of the detected transcript was estimated as 3 kbp approximately.

microdeletion. The six exons (fragments U-Z, see Figure 1) showed no disease related sequence alterations but several polymorphisms. RT-PCR experiments have now revealed that these exons belong to the gene presented here. By using an antisense primer from exon W (Fig. 1) in combination with a sense primer from position 472–492 of the consensus cDNA (Fig. 2), a 1.5 kbp fragment was amplified, and sequence analysis from both ends of the PCR product confirmed the identity of the exons. In total, 2.7 kbp of cDNA sequence were identified, representing at least 18 exons of one gene.

Detection of point mutations in patients with X-linked RP

Mutation screening was carried out in 28 XLRP patients by means of SSCP analysis. Intron primers were designed for the PCR amplification of 10 exons (see Table 1). Five bandshifts were observed in patients but not in controls, indicating disease-specific sequence alterations. The corresponding PCR fragments were sequenced which revealed three different nucleotide exchanges and one 4 bp deletion in five patients (see Fig. 4). Patient 4767 carried a T to G transversion at nucleotide position 420, resulting in the substitution of phenylalanine by cysteine. A proline to serine exchange was shown to be the consequence of a C to T transition at position 734 of the cDNA sequence in patient 2884. Another serine residue was generated due to a G to A exchange at postion 854 in patients 2603 and 2555. Strikingly, the two mutations that give rise to a serine residue are located in a highly conserved domain of the protein which shows 40% identity in a 57 amino acid segment of the RCC1 consensus sequence (RCC1_2.PRF, derived from release 10.0 of PROSITE, see underlined residues in Fig. 4). Moreover, the amino acid substitution in patients 2555 and 2603 affect a glycine residue which is highly conserved during evolution (Fig. 4, position 331). A 4 bp deletion was observed in patient 2550 (nucleotides 1433–1436), resulting in a truncated protein with six abnormal C-terminal amino acids, including a conserved glutamic acid residue (see Figs 4 and 5). None of these changes was detected in 84 male controls.

Table 1. Oligonucleotide primers used for PCR amplification, SSCP screening, and sequencing of exon fragments from the *RP3* gene

Primer ID	Sequence (5′→3′)	Fragment size (bp)	cDNA contained
210	gta gtt ctc ata gta ttc tta cca	235	716–809
211	tta ggt ctt ccc aat cag ctc		
212	ttc aga cac cag ctg ggg tg	204	651–725
213	caa atg tat taa gtg atc ctt gtc		
214	caa ctt tta gga aag ttt ata tca at	289	1276–1456
215	aag gca cat tta tcc tga gag c		
218	gtg tct gct tcc ata atc ctt g	259	810–963
219	tgt aat aaa ata tac cca gtt cta ta		
220	ctg gaa tga gac ctc agt tct c	190	276–338
221	cat ttg tcc tgg act act gtt ca		
222	gga ctc tat agt cta ttg acg tt	238	339–500
223	agg gaa tgt gtc cca gac tga a		
224	gct tca gag cct ggc tac ct	256	501–650
225	aca aca tag aag tgg gag ata ac		
226	ata gat cca tac aag taa cac ttt	250	964–1092
227	tcc atg aaa tac tga aag act ca		
228	cag tgc tag ata ata cta tta tac	273	1093–1275
229	tga agc caa tgt tga tga gtt t		
217	gtt att tta ata aca ggc aac ata g	266	59–180
230	aca cag cag cat atc tat aac aa		
231	tga cat taa aga act aca cag tca	229	181–275
232	tga cct cat ctt aca tta tgt gaa		

In addition to these apparently disease-specific changes we have detected two polymorphisms. One was a G to A transition at position 1305 of the cDNA sequence, resulting in a substitution of arginine by leucine. This change was present in two patients with X-linked RP but also in two out of 86 controls and in the sequence of cosmid ICRFB0972. The second polymorphism was detected at position 1322 (A to G, isoleucine to serine), present in one patient and in five out of 86 controls. Both exchanges represent conservative amino acid substitutions and involve a domain of the predicted gene product that is not homologous to RCC1 proteins.

DISCUSSION

RP3, the gene for the most frequent form of XLRP, had long been known to map to an interval of less than 1000 kbp between the DXS1110 marker and the OTC gene at Xp11.4-p21.1 (3–6,11) but numerous efforts to isolate this gene by positional cloning had remained unsuccessful. By making use of a novel 'YAC Representation Hybridization' approach, to screen this interval for microdeletions, we have recently detected a 6.4 kbp deletion in a patient with XLRP (9). Molecular characterization and sequencing of cosmids and cDNAs spanning this deletion enabled us to define the last six exons of a gene that was expressed in various tissues including retina, and this cDNA was also contained in a larger deletion described by others (12). Here we describe the identification of six novel retina cDNAs and 12 corresponding exons which could be linked to the previously isolated cDNA by RT-PCR and represent the middle and 5′ portion of the same gene. This gene encodes a protein of at least 518 amino acids; the outermost sequences of the ORF have not yet been identified. The predicted protein shows significant homology to the guanine nucleotide exchange factor RCC1 of several species. In five out of 28 patients with XLRP, small sequence alterations were found which were not observed in 84 control males and are therefore considered as disease-specific mutations. In two apparently unrelated patients, the same mutation was observed.

So far, mutations in the *RP3* gene have only been detected in 18% of the patients with XLRP, which is at variance with linkage studies suggesting that *RP3* accounts for at least two-thirds of the familial cases. There are several possible explanations for this discrepancy. First, systematic mutation screening has only been done so far in 16 out of 18 exons identified. Secondly, we have not yet found all exons of the *RP3* gene; at the 5′ end, at least one exon is still missing, and the promotor region of this gene, which has not yet been cloned,

```
          1                                                  50
RCC_HUMAN ...MSPKRIA KRRSPPADAI PK.S..KK..  ...VK....V SHRSHSTEPG
RCC_MESAU ...MPPKRIA KRRSPPEDAI PK.S..KK..  ...VK....V SHRSHKTEPG
RCC_XENLA ...MKGKKTL KRTIAAEESN GT.SDVKK..  ...TKALPIV THPSHGTVGG
RCC_DROME ...MPRRKAL TNNNNAGRAE QQPPKAKR..  ...ARIAPHL ELPKRRTVLG
RCC_YEAST ....MVKRTV ATNGDASGAH RAKKMSKTHA SHIINAQEDY KHMYLSVQPL
RCC_XLRP3 EELMPDSGAV FTFGKSKFAE NNPGKFWFKN DVPVHLSCGD EESAVVTGNN

          51                                                 100
RCC_HUMAN LVLTLGQGDV GQLGLG..EN VMERKKPALV SIPE....DV VQAEAGGMHT
RCC_MESAU LVLTLGQGDV GQLGLG..ES VLERKKPALV PLLQ....DV VQAEAGGMHT
RCC_XENLA QVLTLGQGDV GQLGLG..ED IMERKKPALV TLTK....DI VQAAAGGMHT
RCC_DROME NVLVCGNGDV GQLGLG..ED ILERKRLSPV AGIP....DA VDISAGGMHN
RCC_YEAST DIFCWGTGSM CELGLGPLAK NKEVKRPRLN PFLPRDEAKI ISFAVGGMHT
RCC_XLRP3 KLYMFGSNNW GQLGLGSKSA I..SKPTCVK ALKP...EKV KLAACGRNHT

          101                                                150
RCC_HUMAN VCLSKSGQVY SFGCNDEGAL GRDTS.....  ......VEGS EMVPGKVELQ
RCC_MESAU VCLNQSGQVY SFGCNDEGAL GRDTS.....  ......VEGS EMVPGKVELQ
RCC_XENLA VCLGASGSIY TFGCNDEGAL GRDTS.....  ......EEGS EMQPGKVELA
RCC_DROME LVLTKSGDIY SFGCNDEGAL GRDTS.....  ......EDGS ESKPDLIDLP
RCC_YEAST LALDEESNVW SWGCNDVGAL GRDTSNAKEQ LKDMDADDSS DDEDGDLNEL
RCC_XLRP3 LVSTEGGNVY ATGGNNEGQL G.........  .....LGDTE ERNTFHV...
mut 4767                                                     C

          151                                                200
RCC_HUMAN E.........  .........K VVQVSAGDSH TAALTDDGRV FLWGSFRDNN
RCC_MESAU E.........  .........K VVQVSAGDSH TAALTEDGRV FLWGSFRDNN
RCC_XENLA E.........  .........K VVQVSAGDSH TAALTEDGRV FVFGSFRDNN
RCC_DROME G.........  .........K ALCISAGDSH SACLLKDGRV FAWGSFRDSH
RCC_YEAST ESTPAKIPRE SFPPLAEGHK VVQLAATDNM SCALFSNGEV YAWGTFRCNE
RCC_XLRP3 ..........  .ISFFTSEHK IKQLSAGSNT SAALTEDGRL FMWGD..NSE

          201                                                250
RCC_HUMAN GVIGLLEP..  .MKKSMVPVQ VQLDVPVVKV ASGNDHLVML TADGDLYTLG
RCC_MESAU GVIGLLEP..  .MKKSMVPVQ VQLDMPVVKV ASGNDHLVML TTDGDLYTLG
RCC_XENLA GVIGLLEP..  .MKKSMVPVQ VQINTPVIKI ASGNDHLVLL TVDGDLYTSG
RCC_DROME GNMGLTID..  .GNK.RTPID LMEGTVCCSI ASGADHLVIL TTAGKVFTVG
RCC_YEAST GILGFYQDKI KIQKTPWKVP TFSKYNIVQL APGKDHILFL DEEGMVFAWG
RCC_XLRP3 GQIGLKN...  .VSNVCVPQQ VTIGKPVSWI SCGYYHSAFV TTDGELYVFG

          251                                                300
RCC_HUMAN CGEQGQLGRV PELFANRGGR QGLERL....  LVPKCVMLKS RGSRGHVRFQ
RCC_MESAU CGEQGQLGRV PELFANRGGR QGLERL....  LVPKCVLLKS RGSRGRVRFQ
RCC_XENLA CGEQGQLGRV PERFTNRGGR KGLERL....  LVPQCIHLKA KGS.GRVHFQ
RCC_DROME CAEQGQLGRL SERSISGEGR RGKRDL....  LRPTQLIITR AKP.....FE
RCC_YEAST NGQQNQL...  .........GR KVMERFRLKT LDPRPFGLR.  .....HVKYI
RCC_XLRP3 EPENGKLGLP NQLLGNHRTP QLVSEI....  ..PEKVIQVA CGGEHTVVLT
mut 2884                          S

          301                                                350
RCC_HUMAN DAFCGAYFTF A.ISHEGHVY GFGLSNYHQL G..TPGTESC FIPQNLTSFK
RCC_MESAU DAFCGAYLTF A.ISREGHVY GFGLSNYHQL G..TPGTASC FIPQNLTSFK
RCC_XENLA DVFCGAYFTF A.VSQEGHVY GFGLSNYHQL G..TKNTQAC YAPQNLTSFK
RCC_DROME AIWATNYCTF MRESQTQVIW ATGLNNFKQL A..HETKGKE FA...LTPIK
RCC_YEAST ASGENHCFAL ...TKDNKLV SWGLNQFGQC GVSEDVEDGA LVTKPKRLAL
RCC_XLRP3 ..........  .....KNAVY TFGLGQFGQL GLGTFLFETS .EPKVIENIR
mut 2555/2603                               S

          351                                                400
RCC_HUMAN NSTKSWVGFS GGQHHTVCMD SEGKAYSLGR ..........  ...AEYGRLG
RCC_MESAU NSTKSWVGFS GGQHHTICMD SEGKAYSLGR ..........  ...AEYGRLG
RCC_XENLA NSTKSWIGFS GGQHHTVCVD SEGKAYSLGR ..........  ...AEYGRLG
RCC_DROME TELKDIRHIA GGQHHTVILT TDLKCSVVGR ..........  ...PEYGRLG
RCC_YEAST PDNVVIRSIA AGEHHSLILS QDGDLYSCGR LDMFEVGIPK DNLPEYTYKD
RCC_XLRP3 DQTISY..IS CGENHTALIT DIGLMYTFGD ..........  ...GRHGKLG

          401                                                450
RCC_HUMAN LG.EGAEEKS IPTLISR.LP .AVSSVACGA S.........  ..........
RCC_MESAU LG.EGAEEKS IPTLISR.LP .VVSSVACGA S.........  ..........
RCC_XENLA LG.ENAEEQS EPTPIPD.LP .KINSVASGA S.........  ..........
RCC_DROME LG.DVKDVVE KPTIVKK.LT EKIVSVGCGE V.........  ..........
RCC_YEAST VH.GKARAVP LPTKLNN.VP .KFKSVAAGS H.........  ..........
RCC_XLRP3 LGLENFTNHF IPTLCSNFLR FIVKLVACGG CHMVVFAAPH RGVAKEIEFD

          451                                                500
RCC_HUMAN ..........  ..VGYAVTKD GRVFAWGMGT NYQLGTGQ.D EDAWS.....
RCC_MESAU ..........  ..VGYAVSKD GRVFAWGMGT NYQLGTGQ.D EDAWS.....
RCC_XENLA ..........  ..VSYAVSTD GCVFAWGMGT NLQLGTGE.E EDVWS.....
RCC_DROME ..........  ..CSYAVTID GKLYSWGSGV NNQLGVGD.G DDELE.....
RCC_YEAST ..........  ..HSVAVAQN GIAYSWGFGE TYAVGLGPFE DDTEV.....
RCC_XLRP3 EINDTCLSVA TFLPYSSLTS GNVLQRTLSA RMRRRERERS PDSPSMRRTL

          501                                                550
RCC_HUMAN ....PVEMMG KQLENRVVLS VSSGGQHTVL LVKDKEQS..  ..........
RCC_MESAU ....PVEMTG KQLENRVVLT VSSGGQHTVL LVKDKEQS..  ..........
RCC_XENLA ....PEQMTG KHLEDREVLS VSSGGQHTVL LVRKRS....  ..........
RCC_DROME ....PIVVVS KNTQGKHMLL ASGGGQHAIF LVKADKQDQK  ENVPVKVSGS
RCC_YEAST ....PTRIKN TATQDHNIIL VGCGGQFSVS GGVKLSDEDA  EKRADEMDD.
RCC_XLRP3 PPIEGTLGLS ACFLPNSVFP RCSERNLQES VLSEQDLMQP  EEPDYLLDEM
mut 2550                                          R  NQIIC*

          551                                                600
RCC_HUMAN ..........  ..........  ..........  ..........  ..........
RCC_MESAU ..........  ..........  ..........  ..........  ..........
RCC_XENLA ..........  ..........  ..........  ..........  ..........
RCC_DROME SSISKKDKTP PQDNVDKEAE NVDKQEQKEN LPAKASTSSK KNKTPPQDNA
RCC_YEAST ..........  ..........  ..........  ..........  ..........
RCC_XLRP3 TKEARIDNSS TVESLGETTD ILNMTHIMSL  ..........  ..........
```

might also harbour several of the 'missing' mutations. Mutations in the *RP2* gene, located in the Xp11.2-p11.4 region proximal to *RP3*, is thought to be involved in almost one-third of the cases—but then, the contribution of *RP2* and *RP3* mutations in XLRP is not well defined and may vary in different populations. Finally, the possibility cannot be ruled out that there is yet another XLRP gene in the vicinity of *RP3*.

The predicted amino acid sequence of the *RP3* gene identified in this study shows the highest homology (31.5%) with the human RCC1 protein (13), which acts as a guanine-nucleotide-exchange factor (GEF) of Ran, a GTPase which recently was shown to be involved in nuclear protein import (14). GEFs associate with GDP-bound forms of GTPases and accelerate GDP dissociation and GTP binding, thereby activating the GTPase. A large number of GEFs have been identified, most of which are involved in Ras signalling (15). RCC1 is composed of seven repeats of about 60 residues each. The *RP3* gene thus far is the only gene showing homology to RCC1. In another X-linked eye disorder, choroideremia, the defective gene plays a role in the geranylgeranylation of several different Rab proteins, another family of Ras-related GTPases (7,16). In view of the clinical similarities between choroideremia and *RP3*, it is tempting to speculate that the RP3 protein might function as a GEF for retina specific Rab proteins. Thus far, two different rabGEFs, i.e. Mss4 (mammalian suppressor of sec4) and Dss4–1 (dominant suppressor of sec4) have been cloned (17,18). These proteins do not show any sequence similarity to RCC1 or other GEFs, nor to the RP3 protein described here. Given the large number of GEF molecules involved in ras signalling, the RP3 protein might represent the first member of a new family of GEFs regulating the activity of rabs or related GTPases.

Previously, the presymptomatic diagnosis and carrier detection in XLRP families has been severely hampered by the genetic heterogeneity of this disorder and particularly by the fact that neither clinically nor genetically, RP3 and RP2 could be reliably distinguished. Therefore, the direct detection of *RP3* mutations in presymptomatic affected males and female carriers will also facilitate the diagnosis in other forms of XLRP, and the molecular elucidation of the fundamental defect underlying RP3 may be a clue to the etiology and pathogenesis of RP2, too.

MATERIALS AND METHODS

Construction of a cosmid library

In order to construct a complete cosmid contig from the *RP3* critical region we have subcloned the YAC ICRFy900E0701 into the sCOGH2, an improved version of the recently described cosmid vector sCOGH1 (19). The cosmid library was prepared as described previously (8). Briefly, YAC DNA was partially digested with the restriction enzyme *Sau*3A, separated on a

Figure 4. Sequence alignment of RCC1 proteins from different species by using the PILEUP program. Peptide sequences for the human, hamster, *Xenopus*, *Drosophila* and yeast RCC1 molecules are derived from the Swiss.prot database (accession numbers P18754, P23800, P25183, P25171, P21827). Amino acid residues conserved between the RCC1 polypeptides and the predicted *RP3* gene product are depicted in bold. The RCC1-profile (RCC1_2.PRF, derived from release 10.0 of PROSITE) is underlined. Changes in the amino acid sequence, detected in patients only, are indicated by: mut 4767 (position 145), mut 2884 (position 270), mut 2555/2603 (position 331) and mut 2550 (positions 540–546)

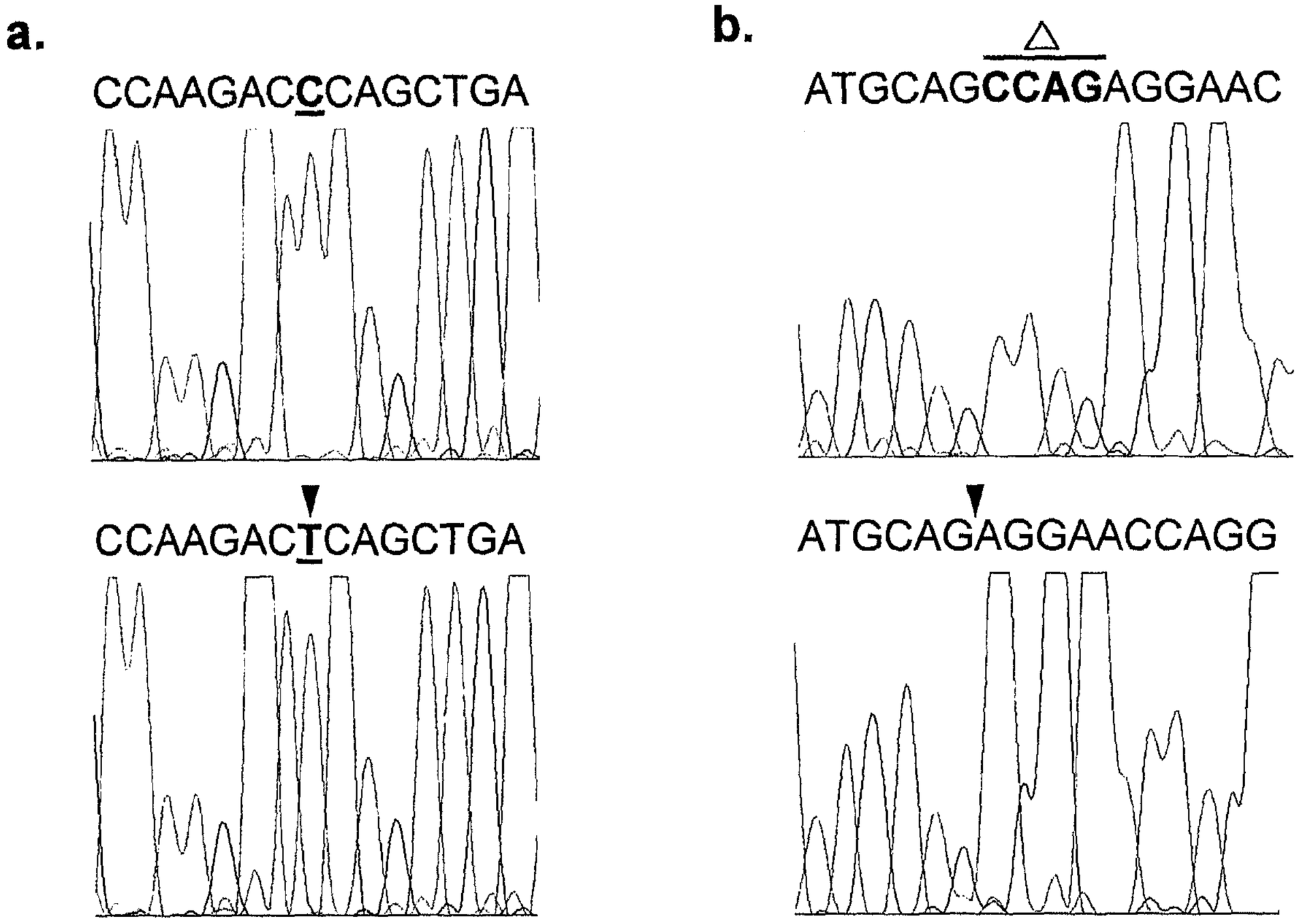

Figure 5. Sequencing patterns from PCR products of controls (top) and patients (bottom). The PCR amplification was performed with primer pairs (**a**) 218 & 219 and (**b**) 214 & 215, and sequencing reactions were carried out with primers 218 and 214, respectively.

sucrose gradient, and the fraction between 30–40 kbp was cloned into the *Bam*HI digested sCOGH2 vector. The ligation mix was packaged by using Stratagene packaging extract (Gold), and 4200 resulting colonies were screened with human placenta DNA. One hundred and forty-four clones containing human DNA inserts were picked and gridded on Nylon membrane (GeneScreenPlus). The contig between the previously identified *RP3* candidate gene (9) and OTC was completed by hybridizing these grids with cosmids LL242D12 and ICRFB0972 as well as the OTC cDNA probe.

Screening of cDNA libraries

The *Not*I insert of cosmid cT-A5 was used to screen a cDNA library, established from adult retina (courtesy of J. Nathans). Screening of 2 million p.f.u. of each library, plated on *E.coli* LE392, was performed as described previously (8). Positive plaques were purified and phage DNA was prepared by using the Qiagen Lambda kit. Inserts were isolated preparatively and subcloned into the pGEM3 vector.

DNA sequencing and data processing

Shotgun cosmid sequencing was performed by dye-primer chemistry on an ABI377 automated sequencer as described previously (9). Assembly of the data was achieved by using the Staden package. Sequencing of the two genomic *Eco*RI fragments (8.0 and 5.8 kbp) and the cDNA clones was carried out by primer walking on subcloned fragments using the DyeDeoxy terminator cycle sequencing kit (ABI). The data were assembled by the GELMERGE program. Exon-intron boundaries were determined by aligning the cDNA sequences to the genomic sequence contig by using BESTFIT and FASTA programs. Prediction of putative exon sequences was done by GRAIL, using releases 2.0 and 1.3 from the GRAIL e-mail server (grail@ornl.gov).

mRNA isolation and RT-PCR

Total RNA was isolated from Hela cells as described elsewhere (20) and poly A^+ mRNA was isolated using the oligotex mRNA kit (QIAgen). RT-PCR was performed mainly as described elsewhere (20), using the sense primer 57 (5′-TGGATCTAATACTTCAGCTGC) and the antisense primer 1136 (5′-GTTTCAGCTGAGCTATCATCA) for PCR amplification of the cDNA products. 35 cycles of denaturation at 95°C for 1 min, annealing at 58°C for 2 min and extension at 72°C for 3 min were carried out with an initial denaturation step of 5 min and a final extension step of 6 min. The PCR products were analysed on a 1.5% agarose gel, purified using the QIAquick gel extraction kit (QIAgen) and sequenced as described above.

Single strand conformation polymorphism (SSCP) analysis

Exon fragments from patients and controls were amplified with intron primers as described (9). The primer sequences, the

corresponding positions within the cDNA and sizes of the fragments are summarized in Table 1. SSCP-PCR was carried out in the presence of $\alpha[^{32}P]$dCTP, and the denatured radiolabeled fragments were separated in a non denaturing polyacrylamide gel as described (21). Sequencing of aberrant PCR fragments was performed by dyedeoxy termination cycle sequencing (ABI) on an ABI370A automated sequencer.

ACKNOWLEDGEMENTS

The authors are grateful to the patients and their families for contributing to this study and making available their blood for DNA isolation. We would like to acknowledge the efforts of S. van der Velde-Visser and L. Boender-van Rossum for establishing immortalized cell lines from the patients and their carrier mothers, and Sven Klages and Michaela Seeger for performing the vast majority of the sequencing reactions. We thank J. den Dunnen for kindly providing the sCOGH2 cosmid vector and J. Nathans for the cDNA library. This work was supported by the Deutsche Forschungsgemeinschaft (G.v.D., Ro 389/16–3) and the Foundation Fighting Blindness, Inc., USA (R. Roepman). F.P.M.C. is recipient of a fellowship of the Royal Dutch Academy of Arts and Sciences.

REFERENCES

1. Jay,M. (1982) On the heredity of retinitis pigmentosa. *Br. J. Ophthalmol.*, **66**, 405–416.
2. Haim,M. (1993) Retinitis pigmentosa: problems associated with genetic classification. *Clin. Genet.*, **44**, 62–70.
3. Teague,P.W., Aldred,M.A., Jay,M., Dempster,M., Harrison,C., Carothers,A.D., Hardwick,L.J., Evans,H.J., Strain,L., Brock,D.J.H., Bundey,S., Jay,B., Bird,A.C., Bhattacharya,S.S. and Wright,A.F. (1994) Heterogeneity analysis in 40 X-linked retinitis pigmentosa families. *Am. J. Hum. Genet.*, **55**, 105–111.
4. Ott,J., Bhattacharya,S., Chen,J.D., Denton,M.J., Donald,J., Dubay,C., Farrar,G.J., Fishman,G.A., Frey,D., Gal,A., Humphries,P., Jay,L., Litt,M., Mächler,M., Musarella,M., Neugebauer,M., Nussbaum,R.L.,Terwilliger,J.D., Weleber,R.G., Wirth,B., Wong,F., Worton,R.G., and Wright,A.F. (1990) Localizing multiple X chromosome linked retinitis pigmentosa loci using multilocus homogeneity tests. *Proc. Natl. Acad. Sci. USA*, **87**, 701–704
5. Chen,J.D., Halliday,F., Keith,G., Sheffield,L., Dickinson,P., Gray,R., Constable,I. and Denton,M. (1989) Linkage heterogeneity between X-linked retinitis pigmentosa and a map of 10 RFLP loci. *Am. J. Hum. Genet.*, **45**, 401–411.
6. Roux,A.-F., Rommens,J., McDowell,C., Anson-Cartwright,L., Bell,S., Schappert,K., Fishman,G.A. and Musarella,M. (1994) Identification of a gene from Xp21 with similarity to the tctex-1 gene of the murine t complex. *Hum. Mol. Genet.*, **3**, 257–263.
7. Cremers,F.P.M., van de Pol,D.J.R., van Kerkhoff,L.P.M., Wieringa,B. and Ropers,H.H. (1990) Cloning of a gene that is rearranged in patients with choroideremia. *Nature*, **347**, 674–677.
8. Berger,W., Meindl,A., van de Pol,T.J.R., Cremers,F.P.M., Ropers,H.H., Dörner,C., Monaco,A., Bergen,A.A.B., Lebo,R., Warburg,M., Zergollern,L., Lorenz,B., Gal,A., Bleeker-Wagemakers,E.M. and Meitinger,T. (1992) Isolation of a candidate gene for Norrie disease by positional cloning. *Nature Genet.*, **1**, 199–203.
9. Roepman,R., Bauer,D., Rosenberg,T., van Duijnhoven,G., van de Vosse,E., Platzer,M., Rosenthal,A., Ropers,H.-H., Cremers,F.P.M. and Berger,W. (1996) Identification of a gene disrupted by a microdeletion in a patient with X-linked retinitis pigmentosa (XLRP). *Hum. Mol. Genet.*, **5**, 827–833.
10. Program Manual for the Wisconsin Package (1994). Genetics Computer Group. Version 8. Madison, Wisconsin, USA
11. Bhattacharya,S.S., Wright,A.F., Clayton,J.F., Price,W.H., Phillips,C.I., McKeown,C.M.E., Jay,M., Bird,A.C., Pearson,P.L., Southern,E.M. and Evans,H.J. (1984) Close genetic linkage between X-linked retinitis pigmentosa and a restriction fragment length polymorphism identified by recombinant DNA probe L1.28. *Nature*, **309**, 253–255.
12. Meindl,A., Carvalho,M.R.S., Herrmann,K., Lorenz,B., Achatz,H., Apfelstedt-Sylla,E., Wittwer,B., Ross,M. and Meitinger,T. (1995) A gene (SRPX) encoding a sushi-repeat-containing protein is deleted in patients with X-linked retinitis pigmentosa. *Hum. Mol. Genet.*, **4**, 2339–2346.
13. Ohtsubo,M., Kai,R., Furuno,N., Sekiguchi,T., Sekiguchi,M., Hayashida,H., Kuma,K., Miyata,T., Fukushige,S., Murotsu,T., Matsubara,K. and Nishimoto,T. (1987) Isolation and characterization of the active cDNA of the human cell cycle gene (RCC1) involved in the regulation of onset of chromosome condensation. *Genes & Dev.*, **1**, 585–593.
14. Nehrbass,U., Blobel,G. (1996) Role of the nuclear transport factor p10 in nuclear import. *Science*, **272**, 120–122.
15. Boguski,M.S., McCormick,F. (1993) Proteins regulating ras and its relatives. *Nature*, **366**, 643–654.
16. Seabra,M.C., Brown,M.S. and Goldstein,J.L. (1993) Retinal degeneration in choroideremia: deficiency of rab geranylgeranyl transferase. *Science*, **259**, 377–381.
17. Moya,M., Roberts,D. and Novick,P. (1993) DSS4–1 is a dominant suppressor of sec4–8 that encodes a nucleotide exchange protein that aids Sec4p function. *Nature*, **361**, 460–463.
18. Burton,J., Roberts,D., Montaldi,M., Novick,P. and De Camilli,P. (1993) A mammalian guanine-nucleotide-releasing protein enhances function of yeast secretory protein Sec4. *Nature*, **361**, 464–467.
19. Datson,N.A., van de Vosse,E., Dauwerse,H.G., Bout,M., van Ommen,G.-J.B. and den Dunnen,J.T. (1996) Scanning for genes in large genomic regions: cosmid-based exon trapping of multiple exons in a single product. *Nucleic Acids Res.*, **24**, 1105–1111.
20. van Bokhoven,H., van den Hurk,J.A.J.M., Bogerd,L., Philipe,C., Gilgenkrantz,S., de Jong,P., Ropers,H.-H. and Cremers,F.P.M. (1994) Cloning and characterization of the human choroideremia gene. *Hum. Mol. Genet.*, **3**, 1041–1046.
21. Berger,W., van de Pol,D., Warburg,M., Gal,A., Bleeker-Wagemakers,L., de Silva,H., Meindl,A., Meitinger,T., Cremers,F. and Ropers,H.H. (1992) Mutations in the candidate gene for Norrie disease. *Hum. Mol. Genet.*, **1**, 461–465.